Non-Standard Limit Theorems in Number Theory

F. Cellarosi[*], Ya.G. Sinai[*,†]

October 18, 2010

There are many unusual limit theorems in Number Theory which are well-known to experts in the field but not so well-known to probabilists. The purpose of this paper is to discuss some examples of such theorems. They were chosen in order to be close to the field of interest of Yu.V. Prokhorov.

One of the main objects in Number Theory is the so-called Möbius function. It is defined as follows

$$\mu(n) = \begin{cases} 1 & \text{if } n=1; \\ 0 & \text{if } n \text{ is not square-free}; \\ (-1)^k & \text{if } n \text{ is the product of } k \text{ distinct primes}. \end{cases}$$

Throughout the paper, when we write $n = p_1 p_2 \cdots p_k$ we assume that $p_1 < p_2 < \ldots < p_k$ are the first k prime numbers. Many properties of the Möbius function are connected with the Riemann zeta function. For example, while the Prime Number Theorem is equivalent to the fact that

$$\sum_{n \leq N} \mu(n) = o(N),$$

the Riemann Hypothesis is equivalent to

$$\sum_{n \leq N} \mu(n) = O_\varepsilon\big(N^{1/2+\varepsilon}\big)$$

for every $\varepsilon > 0$.

Recently, a conjecture by Sarnak [13] has fostered a great interest towards the connections between the Möbius function and Ergodic Theory, and in particular the works of Furstenberg [7] and Green and Tao [9].

---

[*]Mathematics Department, Princeton University. Princeton, NJ, U.S.A.
[†]Landau Institute of Theoretical Physics, Russian Academy of Sciences. Moscow, Russia.

1   A probabilistic model for square-free numbers

Fix $m > 1$ and introduce the set Ω_m, whose elements have the form $n = \prod_{j=1}^{m} p_j^{\nu_j}$, where $\nu_j \in \{0, 1\}$. Then $\mu(n) = \pm 1$ iff $n \in \Omega_m$ for some m. Define on Ω_m the probability distribution Π_m for which

$$\pi_m(n) = \frac{1}{Z_m}\frac{1}{n} = \frac{1}{Z_m \prod_{j=1}^{m} p_j^{\nu_j}}, \tag{1}$$

In (1) Z_m is the normalizing factor and

$$Z_m = \sum_{\nu_1,\ldots,\nu_m} \frac{1}{\prod_{j=1}^{m} p_j^{\nu_j}} = \prod_{j=1}^{m}\left(1 + \frac{1}{p_j}\right) = \exp\left\{\sum_{j=1}^{m} \ln\left(1 + \frac{1}{p_j}\right)\right\} = \exp\left\{O(1) + \sum_{j=1}^{m} \frac{1}{p_j}\right\}$$

as $m \to \infty$. Denote by $N(t)$ the number of primes which are less or equal than t. The Prime Number Theorem says that $N(t) \sim \frac{t}{\ln t}$ as $t \to \infty$ and a slightly stronger version asserts that

$$N(t) - \frac{t}{\ln t} = O\left(\frac{t}{\ln^2 t}\right). \tag{2}$$

We can write, by summation by parts,

$$\sum_{j=1}^{m} \frac{1}{p_j} = \sum_{t=1}^{p_m} \frac{1}{t}\left(N(t) - N(t-1)\right) = \frac{N(p_m)}{p_m + 1} + \sum_{t=1}^{p_m} \frac{N(t)}{t(t+1)} =$$

$$= \frac{m}{p_m + 1} + \sum_{t=1}^{p_m} N(t)\left(\frac{1}{t^2} + O\left(\frac{1}{t^3}\right)\right) = O(1) + \sum_{t=2}^{p_m} \frac{1}{t \ln t} = O(1) + \ln \ln p_m,$$

i.e. $Z_m \sim O(1)\ln p_m$. A more precise asymptotic follows from Mertens' product formula [11]

$$\lim_{n \to \infty} \ln n \prod_{p \leq n}\left(1 - \frac{1}{p}\right) = e^{-\gamma} \approx 0.561459,$$

where γ is Euler-Mascheroni constant. In fact

$$\frac{1}{\ln n}\prod_{p \leq n}\left(1 + \frac{1}{p}\right) = \frac{\left(\prod_{p \leq n} \frac{1}{1-p^{-2}}\right)^{-1}}{\ln n \prod_{p \leq n}\left(1 - \frac{1}{p}\right)} \longrightarrow \frac{\zeta(2)^{-1}}{e^{-\gamma}} \quad \text{as } n \to \infty.$$

Thus

$$Z_m \sim \frac{e^{\gamma}}{\zeta(2)}\ln p_m. \tag{3}$$

By analogy with Statistical Physics, Z_m is called *partition function*.

It is easy to check that w.r.t. Π_m, the random variables ν_j are independent and

$$\Pi_m\{\nu_j = 0\} = \frac{p_j}{1 + p_j}, \quad \Pi_m\{\nu_j = 1\} = \frac{1}{1 + p_j}, \quad 1 \leq j \leq m.$$

Indeed,

$$\Pi_m\{\nu_j = 0\} = \frac{1}{Z_m} \sum_{\nu_1,\ldots,\nu_{j-1}} \sum_{\nu_{j+1},\ldots,\nu_m} \frac{1}{\prod_{l=1}^{j-1} p_l^{\nu_l} \prod_{l=j+1}^{m} p_l^{\nu_l}} =$$

$$= \frac{\prod_{l=1}^{j-1}\left(1 + \frac{1}{p_l}\right) \prod_{l=j+1}^{m}\left(1 + \frac{1}{p_l}\right)}{\prod_{l=1}^{m}\left(1 + \frac{1}{p_l}\right)} = \frac{p_j}{1 + p_j}.$$

Since

$$n = \prod_{j=1}^{m} p_j^{\nu_j} = \exp\left\{\sum_{j=1}^{m} \nu_j \ln p_j\right\},$$

the statistical properties of n w.r.t. Π_m are determined by the properties of $\sum_{j=1}^{m} \nu_j \ln p_j$, which are sums of independent random variables. However the Central Limit Theorem cannot be applied here because ν_j are not identically distributed. Instead, the following limit theorem is valid.

Theorem 1.1. *Let $\zeta_m = \frac{1}{\ln p_m} \sum_{j=1}^{m} \nu_j \ln p_j$. As $m \to \infty$ the distributions of ζ_m converge weakly to the infinitely divisible distribution whose characteristic function $\varphi(\lambda)$ has the form*

$$\varphi(\lambda) = \exp\left\{\int_0^1 \frac{e^{i\lambda v} - 1}{v} dv\right\}. \tag{4}$$

Proof. The characteristic function φ_m of ζ_m is

$$\varphi_m(\lambda) = \mathbb{E}e^{i\lambda\zeta_m} = \mathbb{E}\exp\left\{\frac{i\lambda}{\ln p_m}\sum_{j=1}^{m} \nu_j \ln p_j\right\} = \prod_{j=1}^{m}\left(\frac{p_j}{1+p_j} + \frac{1}{1+p_j}e^{\frac{i\lambda \ln p_j}{\ln p_m}}\right) =$$

$$= \prod_{j=1}^{m}\left(1 + \frac{1}{1+p_j}\left(e^{\frac{i\lambda \ln p_j}{\ln p_m}} - 1\right)\right) =$$

$$= \exp\left\{\sum_{t=1}^{p_m}(N(t) - N(t-1))\ln\left(1 + \frac{1}{1+t}\left(e^{\frac{i\lambda \ln t}{\ln p_m}} - 1\right)\right)\right\} =$$

$$= \exp\left\{f_m(p_m + 1)N(p_m) - \sum_{t=1}^{p_m} N(t-1)(f_m(t+1) - f_m(t))\right\},$$

by summation by parts, where $f_m(s) = \ln\left(1 + \frac{1}{1+s}\left(e^{\frac{i\lambda \ln s}{\ln p_m}} - 1\right)\right)$. Since f_m is complex-valued, the identity $f_m(t+1) - f_m(t) = f_m'(t+\tau)$ for some $0 < \tau < 1$ does not follow from the mean value theorem and we have to work with the real and imaginary parts separately. Writing $f_m = \Re f_m + i\Im f_m$ we have

$$\Re f_m(s) = \ln\left|1 + \frac{1}{1+s}\left(e^{\frac{i\lambda \ln s}{\ln p_m}} - 1\right)\right| = \frac{1}{2}\ln\left(\frac{s^2 + 2s\cos\left(\frac{\lambda \ln s}{\ln p_m}\right) + 1}{(1+s)^2}\right)$$

3

and (by choosing the principal branch of the natural logarithm)

$$\Im f_m(s) = \arg\left(1 + \frac{1}{1+s}\left(e^{\frac{i\lambda \ln s}{\ln p_m}} - 1\right)\right) = \arctan\left(\frac{\sin\left(\frac{\lambda \ln s}{\ln p_m}\right)}{s + \cos\frac{\lambda \ln s}{\ln p_m}}\right).$$

Now, by applying the mean value theorem twice to $\Re f_m$ and $\Im f_m$ separately, we get

$$\Re f_m(t+1) - \Re f_m(t) = (\Re f_m)'(t + \tau_1) = (\Re f_m)'(t) + \tau_1(\Re f_m)''(t + \tau_1')$$

for some $0 < \tau_1' < \tau_1 < 1$, and

$$\Im f_m(t+1) - \Im(f_m)(t) = (\Im f_m)'(t + \tau_2) = (\Im f_m)'(t) + \tau_2(\Im f_m)''(t + \tau_2')$$

for some $0 < \tau_2' < \tau_2 < 1$. Thus

$$\ln\varphi_m(\lambda) = f_m(p_m+1)N(p_m) - \sum_{t=1}^{p_m} N(t-1)\left(f_m'(t) + \tau_1(\Re f_m)''(t + \tau_1') + \tau_2(\Im f_m)''(t + \tau_2')\right). \quad (5)$$

We claim that the sum involving $f_m'(t)$ gives the main term. In fact, the first term and the other sums in (5) tend to zero as $m \to \infty$ (see Appendix). Thus, the main term comes from the following sum:

$$-\sum_{t=1}^{p_m} N(t-1)f_m'(t) = -\sum_{t=2}^{p_m}\left(\frac{t}{\ln t} + O\left(\frac{t}{\ln^2 t}\right)\right)\frac{1}{1 + \frac{1}{1+t}\left(e^{i\lambda\frac{\ln t}{\ln p_m}} - 1\right)} \cdot$$

$$\cdot\left[-\frac{1}{(t+1)^2}\left(e^{i\lambda\frac{\ln t}{\ln p_m}} - 1\right) + \frac{1}{t(t+1)}e^{i\lambda\frac{\ln t}{\ln p_m}}\frac{i\lambda}{\ln p_m}\right] =$$

$$= \sum_{t=1}^{p_m}\left(\frac{1}{t\ln t} + O\left(\frac{1}{t\ln^2 t}\right)\right)\left(1 + \frac{1 - e^{\frac{i\lambda\ln t}{\ln p_m}}}{t + e^{\frac{i\lambda\ln t}{\ln p_m}}}\right) \cdot$$

$$\cdot\left[\left(e^{\frac{i\lambda\ln t}{\ln p_m}} - 1\right) - \frac{(2t+1)\left(e^{\frac{i\lambda\ln t}{\ln p_m}} - 1\right)}{(t+1)^2} - \frac{i\lambda}{\ln p_m}\frac{t}{t+1}e^{\frac{i\lambda\ln t}{\ln p_m}}\right] \quad (6)$$

By opening the brackets in (6) we obtain twelve sums. Let us look at the first sum and consider the change of variables (which will be used in the Appendix too) $v = v(t) = \frac{\ln t}{\ln p_m}$ for which $dv = v(t+1) - v(t) = v'(t + \tau_3) = v'(t) + \tau_3 v''(t + \tau_3')$ for some $0 < \tau_3' < \tau_3 < 1$. We get

$$\sum_{t=2}^{p_m}\frac{1}{t\ln t}\left(e^{\frac{i\lambda\ln t}{\ln p_m}} - 1\right) = \sum\left(dv + \frac{\tau}{(t+\tau')^2\ln p_m}\right)\frac{e^{i\lambda v} - 1}{v} \longrightarrow \int_0^1\frac{e^{i\lambda v} - 1}{v}dv$$

as $m \to \infty$ since for some $C > 0$

$$\left|\sum_{t=2}^{p_m}\frac{\tau}{(t+\tau')^2\ln p_m}\frac{e^{i\lambda v} - 1}{v}\right| \le \frac{C|\lambda|}{\ln p_m}\sum_{t=2}^{p_m}\frac{1}{t^2} \longrightarrow 0.$$

All the remaining eleven sums coming from (6) tend to zero (see Appendix) and this concludes the proof of Theorem 1.1. $\qquad\square$

4

Notice that

$$\int \frac{\cos(\lambda v) - 1}{v} dv = -\int_{|\lambda|v}^{\infty} \frac{\cos u}{u} du - \ln v \qquad \text{and} \qquad \lim_{x \to 0+} \left(-\int_{x}^{\infty} \frac{\cos u}{u} du - \ln x \right) = \gamma,$$

where γ is the Euler-Mascheroni constant as before. Therefore the improper integral $\int_{0}^{1} \frac{\cos(\lambda v) - 1}{v} dv$ converges to $-\gamma - \int_{|\lambda|}^{\infty} \frac{\cos u}{u} du - \ln|\lambda|$. On the other hand

$$\int \frac{\sin(\lambda v)}{v} dv = \text{sgn}(\lambda) \int_{0}^{|\lambda|v} \frac{\sin u}{u} du \qquad \text{gives} \qquad \int_{0}^{1} \frac{\sin(\lambda v)}{v} dv = \text{sgn}(\lambda) \int_{0}^{|\lambda|} \frac{\sin u}{u} du.$$

This shows that

$$\varphi(\lambda) = \begin{cases} \exp\left\{ -\left(\gamma + \int_{\lambda}^{\infty} \frac{\cos u}{u} du + \ln \lambda \right) + i \int_{0}^{\lambda} \frac{\sin u}{u} du \right\} & \lambda > 0, \\ \\ 1 & \lambda = 0, \\ \\ \exp\left\{ -\left(\gamma + \int_{-\lambda}^{\infty} \frac{\cos u}{u} du + \ln(-\lambda) \right) - i \int_{0}^{-\lambda} \frac{\sin u}{u} du \right\} & \lambda < 0. \end{cases}$$

It is known (see [1]) that $\varphi(\lambda)$ is the characteristic function of the *Dickman-De Bruijn distribution*, with density $e^{-\gamma} \rho(t)$, where $\rho(t)$ is determined by the initial condition

$$\rho(t) = \begin{cases} 0, & t \le 0; \\ 1, & 0 < t \le 1, \end{cases} \tag{7}$$

and the integral equation

$$t\rho(t) = \int_{t-1}^{t} \rho(s) ds, \qquad t \in \mathbb{R}.$$

It also satisfies the delay differential equation

$$t\rho'(t) + \rho(t - 1) = 0$$

for $t \ge 1$ (at $t = 1$ we consider the right derivative) and for every $k = 1, 2, 3, \ldots$ there is an analytic function $\rho_k(t)$ that gives $\rho(t)$ on $k - 1 \le t \le k$. For example, $\rho_1 \equiv 1$, $\rho_2(t) = 1 - \ln t$ and $\rho_3(t) = 1 - \ln t + \int_{2}^{t} \ln(u - 1) \frac{du}{u}$. It is also easy to see that $\rho \in C^k([k, \infty))$ for each k.

5

Among other properties of $\rho(t)$ one can mention that it is log-concave on $[1, \infty)$ and

$$\rho(t) = \exp\left\{-t\left(\ln t + \ln\ln t - 1 + \frac{\ln\ln t}{\ln t} + O\left(\frac{(\ln\ln t)^2}{(\ln t)^2}\right)\right)\right\}$$

as $t \to \infty$. In other words, the limiting density $e^{-\gamma}\rho(t)$ is constant on the interval $(0, 1]$, where it takes the value $e^{-\gamma}$, and decays faster then exponentially on $(1, \infty)$, like Poisson distribution. In particular, all its moments exist.

The Dickman-De Bruijn density ρ first appeared in the theory of *smooth numbers* (i.e. numbers with small prime factors). Let $\Psi(x, y)$ denote the number of integers $\leq x$ whose prime factors are $\leq y$. Dickman [4] showed that $\Psi(x, x^{1/u}) \sim x\rho(u)$ as $x \to \infty$. The range of y such that the asymptotic formula $\Psi(x, y) \sim x\rho(u)$, where $x = y^u$, has been significantly enlarged by De Bruijn [1, 2, 3] ($y \geq \exp((\ln x)^{5/8+\varepsilon})$) and Hildebrand [10] ($y \geq \exp((\ln\ln x)^{5/3+\varepsilon})$). Notice that in our ensemble Ω_m (where each element is weighted, not simply counted) we have $x = p_1 p_2 \cdots p_m$ and $y = p_m$ and thus $y \sim \ln x$. In this regime Erdös [6] showed that $\ln\Psi(x, \ln x) \sim \frac{\ln 4 \ln x}{\ln\ln x}$ as $x \to \infty$ and therefore the asymptotic is no longer given by the function ρ. In other words a *phase transition* occurs in the asymptotic behavior of $\Psi(x, y)$. For a survey on the theoretical and computational aspects of smooth numbers see [8].

It is worth to mention that in many limit theorems in Number Theory there appear limiting densities which are constants on some interval starting at 0. An example can be found in the work of Elkies and McMullen [5] on the distribution of the gaps in the sequence $\{\sqrt{n} \mod 1\}$.

Here is another example from Probability Theory where the Dickman-De Bruijn distribution appears. Let $\{\eta_j\}_{j\geq 1}$ be a sequence of independent random variables such that

$$P\{\eta_k = k\} = \frac{1}{k} \qquad \text{and} \qquad P\{\eta_k = 0\} = 1 - \frac{1}{k},$$

and let $\theta_n = \sum_{j=1}^{n} \eta_j$ then

$$\lim_{n\to\infty} P\{n^{-1}\theta_n < x\} = e^{-\gamma}\int_0^x \rho(t)\mathrm{d}t.$$

Theorem 1.1 has several important corollaries and applications. An immediate consequence of (4) is that

$$\Pi_m\{n \leq p_m^s\} = \sum_{n\leq p_m^s, \, n\in\Omega_m} \pi_m(x) \longrightarrow e^{-\gamma}\int_0^s \rho(t)\mathrm{d}t$$

as $m \to \infty$. For instance, for $s = 2$ we get $e^{-\gamma}(3 - \ln 4) \approx 0.90603$. In other words, despite the fact that the largest element of our ensemble Ω_m is of order m^m, approximately 90% of the "mass" of our probability distribution Π_m is concentrated on numbers less than p_m^2 for large m.

Let us fix $0 < \sigma \leq 1$ and decompose the interval $(0, \sigma)$ onto K equal intervals (δ_k, δ_{k+1}), $\delta_k = \frac{\sigma k}{K}$, $k = 0, \ldots, K-1$. For fixed K, Theorem 1.1 states that

$$\Pi_m \left\{ \delta_k < \frac{\ln n}{\ln p_m} < \delta_{k+1} \right\} \longrightarrow \frac{e^{-\gamma} \sigma}{K} \tag{8}$$

as $m \to \infty$. Let us consider the error term in (8)

$$E_m^{(\sigma)}(k, K) := \Pi_m \left\{ \delta_k < \frac{\ln n}{\ln p_m} < \delta_{k+1} \right\} - \frac{e^{-\gamma} \sigma}{K}.$$

In the rest of this paper we provide some estimates about the error terms $E_m^{(\sigma)}(k, K)$ when K grows with n. We prove the following

Theorem 1.2. *For every $\varepsilon > 0$ and every function $K(m)$ such that $\lim_{m \to \infty} \frac{\ln^3 p_m}{K(m)^2} = c \geq 0$ there exists $m^* = m^*(\varepsilon, K)$ such that the inequalities*

$$-\frac{c\sigma^3}{12\zeta(2)} - \varepsilon \leq \frac{Z_m}{p_m^{\sigma}} \sum_{k=0}^{K(m)-1} p_m^{\delta_k} E_m^{(\sigma)}(k, K(m)) \leq \frac{c\sigma^3}{12\zeta(2)} + \varepsilon \tag{9}$$

hold for every $m \geq m^*$ and every $0 < \sigma \leq 1$.

An important tool in the proof of Theorem 1.2 is given by the counting function

$$M_m(t) = \# \{ n \leq t : n \in \Omega_m \} .$$

This is analogous to the classical quantity

$$M(t) = \# \{ n \leq t : \mu(n) \neq 0 \} ,$$

for which the asymptotic

$$\lim_{t \to \infty} \frac{M(t)}{t} = \frac{1}{\zeta(2)} = \frac{6}{\pi^2} \approx 0.607927.$$

holds (see, e.g., [12]). Even though the ensemble Ω_m is very sparse, its initial segment of length p_m contains all square-free numbers less or equal than p_m. In particular $\lim_{m \to \infty} \frac{M_m(p_m^{\sigma})}{p_m^{\sigma}} = \frac{1}{\zeta(2)}$ for every $0 < \sigma \leq 1$. For $\sigma = 1$ this fact can be rephrased as

$$\lim_{m \to \infty} \frac{1}{p_m} \sum_{n \leq p_m} \mu^2(n) = \frac{1}{\zeta(2)}$$

and can be compared with

$$\lim_{m \to \infty} \frac{1}{\ln p_m} \sum_{n \leq p_m} \frac{\mu^2(n)}{n} = e^{-\gamma},$$

which is a corollary of our Theorem 1.1.

The following Lemma provides some simple estimates that will be used in the proof of Theorem 1.2.

Lemma 1.3.

$$0 < \sum_{k=0}^{K-1} p_m^{\delta_{k+1}} \frac{\sigma}{K} - \frac{p_m^{\sigma} - 1}{\ln p_m} \leq \frac{\sigma^3 p_m^{\sigma} \ln^2 p_m}{12K^2} + \frac{\sigma(p_m^{\sigma} - 1)}{2K} \tag{10}$$

$$-\frac{\sigma^3 p_m^{\sigma} \ln^2 p_m}{12K^2} - \frac{\sigma(p_m^{\sigma} - 1)}{2K} \leq \sum_{k=0}^{K-1} p_m^{\delta_k} \frac{\sigma}{K} - \frac{p_m^{\sigma} - 1}{\ln p_m} < 0 \tag{11}$$

Proof. The right (resp. left) Riemann sum $\sum_{k=0}^{K-1} p_m^{\delta_{k+1}} \frac{\sigma}{K}$ (resp. $\sum_{k=0}^{K-1} p_m^{\delta_k} \frac{\sigma}{K}$) converges as $K \to \infty$ to the integral $\int_0^{\sigma} e^{\delta \ln p_m} \mathrm{d}\delta = \frac{p_m^{\sigma} - 1}{\ln p_m}$. Moreover, since the function $t \mapsto p_m^t$ is increasing, the right (resp. left) sum is strictly bigger (resp. smaller) than the integral. This proves the first inequality in (10) and the second inequality in (11). A classical result from Calculus states that in the absolute value of the error performed by approximating the integral $\int_a^b f(x)\mathrm{d}x$ by the trapezoidal Riemann sum

$$\left(\frac{1}{2} f(x_0) + f(x_1) + f(x_2) + \ldots + f(x_{K-1}) + \frac{1}{2} f(x_K) \right) \Delta x,$$

$x_k = a + k \frac{b-a}{K}$ is bounded by $\frac{M(b-a)^3}{12K^2}$ where $\sup_{a \leq x \leq b} |f''(x)| \leq M$. This implies that the error for the right Riemann sum

$$(f(x_1) + \ldots + f(x_K)) \Delta x$$

is bounded from above by $\frac{M(b-a)^3}{12K^2} + (f(b) - f(a))\frac{b-a}{2K}$ and gives the second inequality of (10) when applied to the function $t \mapsto p_m^t$ over the interval $[0, \sigma]$. On the other hand, the error given by the left Riemann sum

$$(f(x_0) + \ldots + f(x_{K-1})) \Delta x$$

is bounded from below by $-\frac{M(b-a)^3}{12K^2} - (f(b) - f(a))\frac{b-a}{2K}$ and this gives the first inequality in (11). $\square$

Proof of Theorem 1.2.

$$\frac{M_m(p_m^{\sigma})}{p_m^{\sigma}} = \frac{Z_m}{p_m^{\sigma}} \sum_{\substack{n \in \Omega_m \\ n \leq p_m^{\sigma}}} n\pi_m(n) = \frac{Z_m}{p_m^{\sigma}} \sum_{k=0}^{K-1} \sum_{\substack{n \in \Omega_m \\ p_m^{\delta_k} < n \leq p_m^{\delta_{k+1}}}} n\pi_m(n) \leq$$

$$\leq \frac{Z_m}{p_m^{\sigma}} \sum_{k=0}^{K-1} p_m^{\delta_{k+1}} \sum_{\substack{n \in \Omega_m \\ p_m^{\delta_k} < n \leq p_m^{\delta_{k+1}}}} \pi_m(n) = \frac{Z_m}{p_m^{\sigma}} e^{-\gamma} \sum_{k=0}^{K-1} p_m^{\delta_{k+1}} \frac{\sigma}{K} + \frac{Z_m}{p_m^{\sigma}} \sum_{k=0}^{K-1} p_m^{\delta_{k+1}} E_m^{(\sigma)}(k, K)$$

Applying Lemma 1.3 to the right Riemann sum $\sum_{k=0}^{K-1} p_m^{\delta_{k+1}} \frac{\sigma}{K}$ we obtain the estimate

$$\frac{M_m(p_m^{\sigma})}{p_m^{\sigma}} \leq \frac{e^{-\gamma} Z_m}{\ln p_m} \frac{p_m^{\sigma} - 1}{p_m^{\sigma}} + \frac{e^{-\gamma} Z_m}{\ln p_m} \left(\frac{\sigma^3 \ln^3 p_m}{12K^2} + \frac{p_m^{\sigma} - 1}{p_m^{\sigma}} \frac{\sigma \ln p_m}{2K} \right) + \frac{Z_m}{p_m^{\sigma}} \sum_{k=0}^{K-1} p_m^{\delta_{k+1}} E_m^{(\sigma)}(k, K)$$

8

which is true for every m and K. Since, as $m \to \infty$, $\frac{M_m(p_m^\sigma)}{p_m^\sigma} \to \frac{1}{\zeta(2)}$, $\frac{Z_m}{\ln p_m} \to \frac{e^\gamma}{\zeta(2)}$, and by hypothesis $\frac{\ln^3 p_m}{K(m)^2} \to c$ (and thus $\frac{\ln p_m}{K(m)} \to 0$), then for every $\varepsilon > 0$ the inequality

$$\frac{Z_m}{p_m^\sigma} \sum_{k=1}^{K(m)-1} p_m^{\delta_{k+1}} E_m^{(\sigma)}(k, K(m)) \geq -\frac{c\sigma^3}{12\zeta(2)} - \varepsilon$$

holds true for sufficiently large m. By noticing that $p_m^{\delta_{k+1}} = p_m^{\delta_k}\left(1 + (e^{\frac{\sigma \ln p_m}{K(m)}} - 1)\right)$ and $0 \leq (e^{\frac{\sigma \ln p_m}{K(m)}} - 1) \to 0$ as $m \to \infty$, we obtain the first inequality of (9). On the other hand

$$\frac{M_m(p_m^\sigma)}{p_m^\sigma} \geq \frac{Z_m}{p_m^\sigma} \sum_{k=0}^{K-1} p_m^{\delta_k} \sum_{\substack{n \in \Omega_m \\ p_m^{\delta_k} < n \leq p_m^{\delta_{k+1}}}} \pi_m(n) = \frac{Z_m}{p_m^\sigma} e^{-\gamma} \sum_{k=0}^{K-1} p_m^{\delta_k} \frac{\sigma}{K} + \frac{Z_m}{p_m^\sigma} \sum_{k=0}^{K-1} p_m^{\delta_k} E_m^{(\sigma)}(k, K)$$

and applying Lemma 1.3 to the left Riemann sum $\sum_{k=0}^{K-1} p_m^{\delta_k} \frac{\sigma}{K}$ we obtain the estimate

$$\frac{M_m(p_m^\sigma)}{p_m^\sigma} \geq \frac{e^{-\gamma} Z_m}{\ln p_m} \frac{p_m^\sigma - 1}{p_m^\sigma} - \frac{e^{-\gamma} Z_m}{\ln p_m} \left(\frac{\sigma^3 \ln^3 p_m}{12K^2} + \frac{p_m^\sigma - 1}{p_m^\sigma} \frac{\sigma \ln p_m}{2K} \right) + \frac{Z_m}{p_m^\sigma} \sum_{k=0}^{K-1} p_m^{\delta_k} E_m^{(\sigma)}(k, K)$$

which is true for every m and K. Proceeding as above we have that for every $\varepsilon > 0$ the inequality

$$\frac{Z_m}{p_m^\sigma} \sum_{k=0}^{K(m)-1} p_m^{\delta_k} E_m^{(\sigma)}(k, K(m)) \leq \frac{c\sigma^3}{12\zeta(2)} + \varepsilon$$

holds for sufficiently large m and we have the second inequality of (9). $\qquad \square$

An immediate consequence of Theorem 1.2 is the following

Corollary 1.4. *Consider a function $K(m)$ such that $\lim_{m \to \infty} \frac{\ln^3 p_m}{K(m)^2} = c \geq 0$. Then the sum of the error terms coming from (8), with weights $p_m^{-\sigma+\delta_k}$, satisfies the asymptotic estimate*

$$\sum_{k=0}^{K(m)-1} \frac{E_m^{(\sigma)}(k, K(m))}{p_m^{\sigma-\delta_k}} = \begin{cases} O\left(\frac{1}{\ln p_m}\right) & \text{if } c > 0; \\ o\left(\frac{1}{\ln p_m}\right) & \text{if } c = 0; \end{cases} \tag{12}$$

as $m \to \infty$ for every $0 < \sigma \leq 1$.

Notice that implied constant in the O-notation depends explicitly on c and σ by (3) and (9). Moreover, as k ranges from 0 to $K(m) - 1$, the weights vary from $p_m^{-\sigma}$ ($\to 0$ as $m \to \infty$) to $e^{-\frac{\ln p_m}{K(m)}}$ ($\to 1$ as $m \to \infty$). This means that the error terms $E_m^{(\sigma)}(k, K(m))$ corresponding to small values of k are allowed to be larger in absolute value.

9

In order to get estimates on the mean value of the error term (for which al weights are equal to $\frac{1}{K(m)}$) we just replace the weights $p_m^{-\sigma+\delta_k}$ by either $p^{-\sigma}$ or 1 in (9). This yields, for every ε and sufficiently large m,

$$\frac{1}{Z_m}\left(-\frac{c\sigma^3}{12\zeta(2)}-\varepsilon\right) \le \sum_{k=0}^{K(m)-1} E_m^{(\sigma)}(k, K(m)) \le \frac{p_m^\sigma}{Z_m}\left(\frac{c\sigma^3}{12\zeta(2)}+\varepsilon\right).$$

In particular we get, as $m \to \infty$,

$$\langle E_m^{(\sigma)}\rangle := \frac{1}{K(m)}\sum_{k=0}^{K(m)-1} E_m^{(\sigma)}(k, K(m)) = \begin{cases} O\left(\frac{p_m^\sigma}{\ln^{5/2} p_m}\right) & \text{if } c > 0; \\[3mm] o\left(\frac{p_m^\sigma}{K(m)\ln p_m}\right) & \text{if } c = 0. \end{cases}$$

Let us point out that, even though by (8) the error term $E_m^{(\sigma)}(k, K(m))$ tends to zero as $m \to \infty$ for each k, it is not a priori true that $\langle E_m^{(\sigma)}\rangle$ tends to zero as well. It follows from our Theorem 1.2 that this is indeed the case when $\frac{p_m^\sigma}{K(m)\ln p_m}$ remains bounded (i.e. a particular case of $c = 0$). Let us summarize this fact in the following

Corollary 1.5. *Let $0 < \sigma \le 1$ and consider a function $K(m)$ such that $\lim_{m\to\infty} \frac{p_m^\sigma}{K(m)\ln p_m} < \infty$. Then, as $m \to \infty$,*

$$\langle E_m^{(\sigma)}\rangle = o\left(\frac{p_m^\sigma}{K(m)\ln p_m}\right). \tag{13}$$

In other words, if K grows sufficiently fast (namely as $const \cdot \frac{p_m^\sigma}{\ln p_m}$ or faster), then the mean value of the error $\langle E_m^{(\sigma)}\rangle$ tends to zero as $m \to \infty$ and the rate of convergence to zero is controlled explicitly in terms of σ and K.

Notice that one would expect the error term $E_m^{(\sigma)}(k, K(m))$ in (8) to be $o\left(\frac{1}{K(m)}\right)$, however we could only derive the weaker asymptotic estimates (12) and (13) from Theorem 1.2. A possible approach to further investigate the size of the error term in (8) would be to first prove an analogue of Theorem 1.1 for shrinking intervals. This is, however, beyond the aim of this paper.

Appendix

This Appendix contains the estimates for the error terms in the proof Theorem 1.1. By C_j, $j = 1, \ldots, 21$, we will denote some positive constants.

The first term of (5) tends to zero as $m \to \infty$ uniformly in λ. In fact using (2) we obtain

$$\Re f_m(p_m+1)N(p_m) = \frac{N(p_m)}{2}\ln\left(\frac{(p_m+1)^2 + 2(p_m+1)\cos\left(\frac{\lambda\ln(p_m+1)}{\ln p_m}\right) + 1}{(p_m+2)^2}\right) =$$

$$= \frac{N(p_m)}{2}\left(\ln\left(1+O\left(\frac{1}{p_m}\right)\right) - \ln\left(1+O\left(\frac{1}{p_m}\right)\right)\right) = O\left(\frac{p_m}{\ln p_m}\right)O\left(\frac{1}{p_m}\right) = O\left(\frac{1}{\ln p_m}\right),$$

and

$$\Im f_m(p_m+1)N(p_m) = N(p_m)\arctan\left(\frac{\sin\left(\lambda\frac{\ln(p_m+1)}{\ln p_m}\right)}{p_m+1+\cos\left(\lambda\frac{\ln(p_m+1)}{\ln p_m}\right)}\right) =$$

$$= O\left(\frac{p_m}{\ln p_m}\right)O\left(\frac{1}{p_m}\right) = O\left(\frac{1}{\ln p_m}\right)$$

as $m \to \infty$, and the implied constants do not depend on λ. An explicit computation shows that

$$(\Re f_m)''(s) = f_m^{(1)}(s) + f_m^{(2)}(s) + f_m^{(3)}(s),$$

where

$$f_m^{(1)}(s) = -\lambda^2 \frac{2s + (1+s^2)\cos\left(\frac{\lambda\ln s}{\ln p_m}\right)}{s\left(s^2 + 2s\cos\left(\frac{\lambda\log s}{\log p_m}\right) + 1\right)^2 \ln^2 p_m},$$

$$f_m^{(2)}(s) = \lambda \frac{\left(3s^2 + 2s\cos\left(\frac{\lambda\ln s}{\ln p_m}\right) - 1\right)\sin\left(\frac{\lambda\ln s}{\ln p_m}\right)}{s\left(s^2 + 2s\cos\left(\frac{\lambda\log s}{\log p_m}\right) + 1\right)^2 \ln p_m},$$

$$f_m^{(3)}(s) = \frac{2\left(\cos\left(\frac{\lambda\ln s}{\ln p_m}\right) - 1\right)\left(s^3 - s^2 - s - 1 + (s^2 - 2s - 1)\cos\left(\frac{\lambda\ln s}{\ln p_m}\right)\right)}{(1+s)^2\left(s^2 + 2s\cos\left(\frac{\lambda\log s}{\log p_m}\right) + 1\right)^2}.$$

We have

$$\left|f_m^{(1)}(s)\right| \le \frac{C_1\lambda^2}{s^3\ln^2 p_m}, \qquad \left|f_m^{(2)}(s)\right| \le \frac{C_2|\lambda|}{s^3\ln p_m}$$

and thus

$$\left|\sum_{t=1}^{p_m} N(t-1)\tau_1 f_m^{(1)}(t+\tau_1')\right| \le \frac{C_3\lambda^2}{\ln^2 p_m}\sum_{t=2}^{p_m}\frac{1}{t^2\ln t} \longrightarrow 0 \quad \text{and}$$

$$\left|\sum_{t=1}^{p_m} N(t-1)\tau_1 f_m^{(2)}(t+\tau_1')\right| \le \frac{C_4|\lambda|}{\ln p_m}\sum_{t=2}^{p_m}\frac{1}{t^2\ln t} \longrightarrow 0 \quad \text{as } m \to \infty.$$

The third function satisfies the estimate

$$\left|f_m^{(3)}(s)\right| \le \frac{s^3\left|2\cos\left(\frac{\lambda\ln s}{\ln p_m}\right) - 2\right| + s^2 C_5}{(1+s)^2(1-s)^4} \le \frac{C_6\left(1 - \cos\left(\frac{\lambda\ln s}{\ln p_m}\right)\right)}{s^3}.$$

We now perform the same change of variables $v = v(t) = \frac{\ln t}{\ln p_m}$ as before (using τ_3 and τ_3' as in the proof of Theorem 1.1). We get

$$\left|\sum_{t=1}^{p_m} N(t-1)\tau_1 f_m^{(3)}(t+\tau_1')\right| \le C_7 \sum_{t=2}^{p_m} \frac{1 - \cos\left(\frac{\lambda\ln t}{\ln p_m}\right)}{t^2\ln t} \le$$

$$\le C_8 \sum_v \left(dv + \frac{\tau_3}{(t+\tau_3')^2\ln p_m}\right)\frac{1 - \cos(\lambda u)}{t\,v} \longrightarrow 0$$

11

as $m \to \infty$. Another explicit computation shows that

$$(\Im f_m)''(s) = f_m^{(4)}(s) + f_m^{(5)}(s) + f_m^{(6)}(s),$$

where

$$f_m^{(4)}(s) = \lambda^2 \frac{(s^2 - 1)\sin\left(\frac{\lambda \log s}{\log p_m}\right)}{s\left(s^2 + 2s\cos\left(\frac{\lambda \log s}{\log p_m}\right) + 1\right)^2 \ln^2 p_m},$$

$$f_m^{(5)}(s) = -\lambda \frac{1 + 5s^2 + 2s^2\cos^2\left(\frac{\lambda \log s}{\log p_m}\right) + (3s^3 + 5s)\cos\left(\frac{\lambda \log s}{\log p_m}\right)}{s^2\left(s^2 + 2s\cos\left(\frac{\lambda \log s}{\log p_m}\right) + 1\right)^2 \ln p_m}$$

$$f_m^{(6)}(s) = \frac{2\left(s + \cos\left(\frac{\lambda \log s}{\log p_m}\right)\right)\sin\left(\frac{\lambda \log s}{\log p_m}\right)}{\left(s^2 + 2s\cos\left(\frac{\lambda \log s}{\log p_m}\right) + 1\right)^2}.$$

We have the estimates

$$\left|f_m^{(4)}(s)\right| \le \frac{C_{10}\lambda^2}{s^3 \ln^2 p_m}, \qquad \left|f_m^{(5)}(s)\right| \le \frac{C_{11}|\lambda|}{s^3 \ln p_m}$$

and thus

$$\left|\sum_{t=1}^{p_m} N(t-1)\tau_1 f_m^{(4)}(t + \tau_1')\right| \le \frac{C_{12}\lambda^2}{\ln^2 p_m} \sum_{t=2}^{p_m} \frac{1}{t^2 \ln t} \longrightarrow 0 \quad \text{and}$$

$$\left|\sum_{t=1}^{p_m} N(t-1)\tau_1 f_m^{(5)}(t + \tau_1')\right| \le \frac{C_{13}|\lambda|}{\ln p_m} \sum_{t=2}^{p_m} \frac{1}{t^2 \ln t} \longrightarrow 0 \quad \text{as } m \to \infty.$$

The estimate

$$\left|f_m^{(6)}(s)\right| \le \frac{C_{14}s\sin\left(\frac{\lambda \log s}{\log p_m}\right)}{(s-1)^4} \le \frac{C_{15}\sin\left(\frac{\lambda \log s}{\log p_m}\right)}{s^3}$$

yields, as $m \to \infty$,

$$\left|\sum_{t=1}^{p_m} N(t-1)\tau_1 f_m^{(6)}(t + \tau_1')\right| \le C_{15} \sum_{t=2}^{p_m} \frac{\sin\left(\frac{\lambda \ln t}{\ln p_m}\right)}{t^2 \ln t} \le C_{16} \sum_v \left(dv + \frac{\tau_3}{(t + \tau_3')^2 \ln p_m}\right) \frac{\sin(\lambda v)}{t\, v} \to 0.$$

This concludes the analysis of the error terms coming from (5).

Let us now deal with the error terms coming from (6). One sum (giving the main term) is already discussed in the proof of Theorem 1.1. Amongst the remaining eleven sums coming from (6), it is enough to check that the following three tend to zero as $m \to \infty$ (the other eight being

12

dominated by these):

$$\left| \sum_{t=2}^{p_m} \frac{1}{t \ln t} \frac{(2t-1)\left(e^{\frac{i\lambda \ln t}{\ln p_m}} - 1\right)}{(t+1)^2} \right| \le C_{17} \sum_{t=1}^{p_m} \left| \frac{e^{\frac{i\lambda \ln t}{\ln p_m}} - 1}{\frac{\ln t}{\ln p_m}} \right| \frac{1}{t^2 \ln p_m} \le \frac{C_{18}|\lambda|}{\ln p_m} \sum_{t=2}^{p_m} \frac{1}{t^2} \longrightarrow 0,$$

$$\left| \frac{i\lambda}{\ln m} \sum_{t=2}^{p_m} \frac{1}{t \ln t} \frac{t}{t+1} e^{\frac{i\lambda \ln t}{\ln p_m}} \right| \le \frac{C_{19}|\lambda|}{\ln m} \sum_v \left(\mathrm{d}v + \frac{\tau}{(t+\tau')^2 \ln p_m} \right) \frac{e^{i\lambda v}}{v} \longrightarrow 0,$$

$$\left| \sum_{t+2}^{p_m} \frac{1}{t \ln t} \frac{\left(e^{\frac{i\lambda \ln t}{\ln p_m}} - 1\right)^2}{t + e^{\frac{i\lambda \ln t}{\ln p_m}}} \right| \le C_{20} \sum_{t=2}^{p_m} \left| \frac{\left(e^{\frac{i\lambda \ln t}{\ln p_m}} - 1\right)^2}{\left(\frac{\ln t}{\ln p_m}\right)^2} \right| \frac{\ln t}{t^2 \ln^2 p_m} \le \frac{C_{21}\lambda^2}{\ln^2 p_m} \sum_{t=2}^{p_m} \frac{\ln t}{t^2} \longrightarrow 0.$$

Acknowledgments

References

[1] N. G. de Bruijn. The asymptotic behaviour of a function occurring in the theory of primes. *J. Indian Math. Soc. (N.S.)*, 15:25–32, 1951.

[2] N. G. de Bruijn. On the number of positive integers $\le x$ and free of prime factors $> y$. *Nederl. Acad. Wetensch. Proc. Ser. A.*, 54:50–60, 1951.

[3] N. G. de Bruijn. On the number of positive integers $\le x$ and free prime factors $> y$. II. *Nederl. Akad. Wetensch. Proc. Ser. A 69=Indag. Math.*, 28:239–247, 1966.

[4] K. Dickman. On the frequency of numbers containing primes of a certain relative magnitude. *Ark. Mat. Astr. Fys. 22*, 22A:1–14, 1930.

[5] N. D. Elkies and C. T. McMullen. Gaps in $\sqrt{n}$ mod 1 and ergodic theory. *Duke Math. J.*, 123(1):95–139, 2004.

[6] P. Erdös. *Wiskundige Opgaven met de Oplossingen*, 21:Problem and Solution Nr. 136, 1963.

[7] H. Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory*, 1:1–49, 1967.

[8] A. Granville. Smooth numbers: computational number theory and beyond. In *Algorithmic number theory: lattices, number fields, curves and cryptography*, volume 44 of *Math. Sci. Res. Inst. Publ.*, pages 267–323. Cambridge Univ. Press, Cambridge, 2008.

[9] B. Green and T. Tao. The mobius function is strongly orthogonal to nilsequences. *To appear in Annals of Math.*

[10] A. Hildebrand. Integers free of large prime factors and the Riemann hypothesis. *Mathematika*, 31(2):258–271 (1985), 1984.

[11] F. Mertens. Ein beitrag zur analytischen zahlentheorie. Ueber die vertheilung der primzahlen. *J. reine angew. Math*, 78:46–62, 1874.

[12] H. L. Montgomery and R. C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.

[13] P Sarnak. Möbius randomness and dynamics. *Lecture Slides Summer 2010.* $http://www.math.princeton.edu/sarnak/Mobius\%20lectures\%20Summer\%202010.pdf$.